A COMPARISON BETWEEN STEEPEST ASCENT AND DIFFERENTIAL CORRECTION OPTIMIZATION METHODS IN A PROBLEM OF BOLZA, WITH A METHOD FOR OBTAINING STARTING VALUES FOR THE ADJOINT VARIABLES FROM A NOMINAL PATH

WILLIAM W. McCUE
and
ROBERT C. GOOD

36

A COMPARISON BETWEEN STEEPEST ASCENT AND DIFFERENTIAL CORRECTION OPTIMIZATION METHODS IN A PROBLEM OF BOLZA, WITH A METHOD FOR OBTAINING STARTING VALUES FOR THE ADJOINT VARIABLES FROM A NOMINAL PATH

by

William W. McCue

Lieutenant Commander, United States Navy

and

Robert C. Good

Lieutenant, United States Navy

Submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE IN MATHEMATICS
(Lieutenant Good)

MASTER OF SCIENCE IN PHYSICS
(Lieutenant Commander McCue)

United States Naval Postgraduate School
Monterey, California

1 9 6 3

A Comparison Between Steepest Ascent and Differential Correction Optimization Methods in a Problem of Bolza, with a Method for Obtaining Starting Values for the Adjoint Variables from a Nominal Path

* * * * *

William W. McCue

and

Robert C. Good

A COMPARISON BETWEEN STEEPEST ASCENT AND DIFFERENTIAL CORRECTION OPTIMIZATION METHODS IN A PROBLEM OF BOLZA, WITH A METHOD FOR OBTAINING STARTING VALUES FOR THE ADJOINT VARIABLES FROM A NOMINAL PATH

by

William W. McCue

and

Robert C. Good

This work is accepted as fulfulling

the thesis requirements for the degree of

MASTER OF SCIENCE IN PHYSICS
(Lieutenant Commander McCue)

MASTER OF SCIENCE IN MATHEMATICS
(Lieutenant Good)

from the

United States Naval Postgraduate School

ABSTRACT

The problem of maximum range atmospheric reentry for an orbiting lifting glider was treated by Bryson and Denham[1] by the method of "steepest ascent". The same problem is undertaken here by a method of differential corrections developed by Faulkner. This method makes use of a Newton-Raphson type iteration based on paths which satisfy the Euler-Lagrange equations. A comparison of results is made, showing large differences in control variable history, and longer range for the path obtained by differential corrections. The problem was characterized by a sharp "ridge" in the domain of the starting values of the adjoint variables and the effect of this on the convergence of both methods is discussed. Finally, the difficulty of choosing initial approximations for the starting values of the adjoint variables is discussed, and a method is presented for obtaining these from a nominal path as the first step in the computer routine.

[1] A. E. Bryson and W. F. Denham, A steepest-ascent method for solving optimum programming problems (Ratheon report BR-1303), Raytheon Company, Bedford, Mass., 1961.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

INTRODUCTION

A problem of optimum control may be described very roughly as the determination of the decision-making or control function which will result in the largest possible value of the final "payoff" from some continuing process.

The process might be illustrated as shown in figure 1.1, where P is the "payoff", something whose value depends on the terminal point or the path history, and Q is a set of given constraints which must be satisfied.

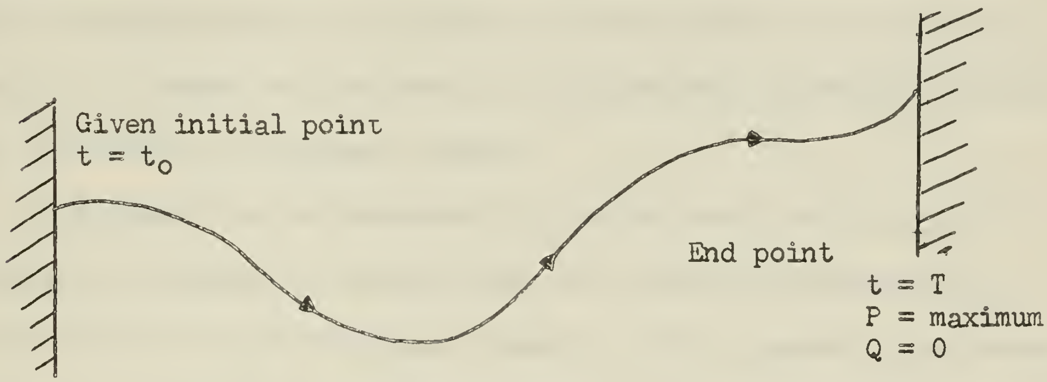


Fig. 1.1 An Optimum Path

The methods of solution commonly used now were foreseen at least forty years ago, but the volume of computation required limited their application to only the simplest of problems. In the past ten years there has been an explosion of interest due to the availability of high-speed digital computers on the one hand, and on the other hand to the urgent need for solutions in such applications as space trajectories. The "state of the art" is such that no method is com-

pletely general, and a great diversity of approaches is employed, reflecting the absence of a firm framework of mathematical theory concerning nonlinear differential equations.

In this thesis a comparison is made between two methods for calculating optimum angle of attack programs for a lifting unpowered vehicle to attain maximum range, starting from circular orbital velocity at a height of 300 thousand feet above the surface of the earth. The problem was chosen because it is typical of a class of problems in which there is now great interest, and because two of the foremost workers in the field of optimization, Arthur E. Bryson and W. F. Denham, had published a clear and well-documented solution by the method of "steepest ascent".

The comparison was suggested by Stanley Ross, while one of the writers was engaged in summer field work under the sponsorship of the Lockheed Missile and Space Company. John V. Breakwell and George Leitmann generously gave many hours of their time to general discussions and made specific suggestions which were very helpful. Professor Frank D. Faulkner of the U. S. Naval Postgraduate School not only laid out the general form of the problem and overcame the difficulties that arose, but also the method of solution used was the one developed by him from the basic method of differentials of Bliss.

Chapter I

GENERAL FORMULATION OF A PROBLEM OF OPTIMUM CONTROL

1.1 The equations of motion, the control variables and the constraints.

It will be helpful to define an optimum control problem in somewhat more exact terms than those given in the introductory remarks.

Suppose we are given a system whose behavior is described by a system of differential equations which we shall call the equations of motion. We will suppose that wherever any of these differential equations is of higher than first order, we have defined such additional variables as necessary so as to obtain n first order differential equations of the form

$$\dot{s}_i = f_i(s_1, s_2, \ldots, s_n, p_1, p_2, \ldots, p_m, t) \qquad i = 1, 2, \ldots, n$$

where the variables s_i are the state variables, the p_j are the control variables, t is the independent variable, and the dot denotes differentiation with respect to the independent variable. It is assumed that the f_i are of class C".

It will be convenient for what is to follow, if matrix notation is adopted at this point. The equations of motion may be written as the single matrix equation

$$(1.1) \qquad \dot{S} = F(S, P, t)$$

where

$$S = \begin{bmatrix} s_1 \\ \cdot \\ \cdot \\ \cdot \\ s_n \end{bmatrix}$$, an nxl matrix of the dependent <u>state variables</u>

$$F = \begin{bmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{bmatrix}$$, an nxl matrix of known functions of the control variables, the state variables, and possibly the independent variable

$$P = \begin{bmatrix} p_1 \\ \cdot \\ \cdot \\ \cdot \\ p_m \end{bmatrix}$$, an mxl matrix of <u>control</u> <u>variables</u>, which we are free to choose as functions of t.

In addition, there are k constraints or end conditions, of the form $e_i(S,t) = 0$, which may define values of the state variables at the end point $t = T$, or may be of integral form such as

$$\int_{t_1}^{t_2} f_1 \, dt = \text{constant}.$$

In the latter case, we shall define an additional variable s_{n+1} so that

$$\dot{s}_{n+1} = f_i$$

with the initial and end conditions

$$s_{n+1}(0) = 0$$

$$s_{n+1}(T) = \text{the given constant.}$$

These may be written as

$$(1.2) \qquad E = \begin{bmatrix} e_1 \\ \cdot \\ \cdot \\ \cdot \\ e_k \end{bmatrix}, \text{ a kx1 matrix of } \underline{\text{constraint}}\ \underline{\text{functions}}.$$

We wish to discover the particular control variable program which will drive this system from a given initial point to some terminal point which, of the set of all such points which satisfy the terminal constraints, gives a maximum value to some function, $M(S,t)$, of the state variables; that is, for the terminal point $t = T$, possibly unknown,

$$M(S,t)_T = M(T) = \max.$$

1.2 Bounded control variables and inequality constraints.

The permissible domain of the control variables may likewise be limited. Some examples of restrictions on the control for common physical systems might be:

$$\sum_1 F_i \leqslant P \max,$$

$$\int_0^{t_1} P^1\, dt \leqslant I \max,$$

$$P = P_{max} \quad \text{for } 0 \leq t \leq t_1 \text{ and}$$

$$P = 0 \quad \text{for } t_1 < t < T.$$

Bounded control variables appear frequently in real physical systems. For solving these problems it may be useful to replace the inequalities by mathematically equivalent equalities. This is accomplished by the introduction of suitable new variables in the manner due to Valentine (ref.1). For instance, for a constraint

$$0 \leq p \leq P\ max$$

one may define the real variable g, so that

$$p(P_{max} - p) - g^2 = 0.$$

This procedure transforms the inequality to an equality, and we adjoin this equation to the set (1.1). The same procedure is applied in the case of inequality constraints or end conditions.

We will generally require that all state variables be specified initially, but this is not necessary. One or more of these may be unspecified, that is "free".

The general form of these problems, then, is the two-point boundary value problem. With the stipulation that the functional M is to be maximized (or minimized) they become, depending on the formulation, the classical problem of Bolza or problem of Mayer of the calculus of variations. Since to maximize some quantity, say u, is equivalent to minimizing the quantity minus u, we will henceforth consider that the term "maximum" includes both cases

unless otherwise specified.

1.3 Degenerate problems.

Some important observations may be made here. First, as pointed out by Faulkner (ref.2) and Breakwell (ref.3), the case in which all of the functions f_i involve any function of a particular control variable only linearly are said to be degenerate. These degenerate problems typically are of the "bang-bang" type in which the control changes discontinuously. The neglect of induced drag in aerodynamic problems for instance often results in bang-bang control.

Second, the constraints must not involve the control variables, or the problem is overspecified. If the control program is to be such as to produce an extreme of the functional M, then it may not simultaneously satisfy any prescribed constraint at any isolated point of the path, else the solution is generally discontinuous.

1.4 The adjoint variables.

We shall introduce a set of Lagrange multipliers, as yet undetermined

$$U = \left[u_1, u_2, \ldots, u_n\right]$$

a lxn matrix of adjoint variables. The term "adjoint" is used, since they will be chosen to be solutions to the system of differential equations which is adjoint to equations for the variations of the given system.

Suppose we write (1.1) as

$$\dot{S} - F = 0$$

and integrate its product with the adjoint vector between the limits of the independent variable. For convenience, we will call these limits zero and T, and we have,

$$\int_0^T U(\dot{S} - F)\, dt = 0.$$

The corresponding variational equation is

$$\int_0^T U(\delta\dot{S} - F_s\delta S - F_p\delta P)\, dt = 0.$$

Now this is integrated by parts to eliminate from the integrand the variations of the state variables.

$$\text{(1.3)} \qquad \Big[U\delta S\Big]_0^T = \int_0^T\Big[(\dot{U} + UF_s)\delta S + U\ F_p\delta P\Big]\, dt - \Big[\sum_k U\ F\Big]_{t_{k-}}^{t_{k+}} dt_k$$

where F_s and F_p are the matrices of partial derivatives,

$$\text{(1.4)} \qquad F_s = \begin{bmatrix} \partial f_1/\partial s_1 & \cdots & \partial f_1/\partial s_n \\ \vdots & & \vdots \\ \partial f_n/\partial s_1 & \cdots & \partial f_n/\partial s_n \end{bmatrix}, \text{ an nxn matrix,}$$

$$F_p = \begin{bmatrix} \partial f_1/\partial p_1 & \cdots & \partial f_1/\partial p_m \\ \vdots & & \\ \partial f_n/\partial p_1 & \cdots & \partial f_1/\partial p_m \end{bmatrix}, \text{ an nxm matrix,}$$

and t_k is a symbol for any point or set of points where $\dot{S}$ or U is discontinuous.

Now, if U is chosen as a solution to the equation

(1.5) $$\dot{U} + U F_s = 0,$$

then (3) becomes

(1.6) $$\left[U\delta S\right]_0^T = \int_0^T (U F_p \delta P)\, dt - \sum_k \left[UF\right]_{t_k^-}^{t_k^+} dt_k .$$

This sequence of operations defines the system which is adjoint to the variational equations of (1.1), $\delta\dot{S} - F_s \delta S = F_p \delta P$, and forms the 1 x n matrix, or vector

$$\dot{U} = \left[\dot{u}_1 \ \dot{u}_2 \ . \ . \ . \ \dot{u}_n\right] .$$

Hereafter, no distinction will be drawn between a vector in k-space and a 1 x k or k x 1 matrix, or a corresponding column or row of a matrix.

1.5 Green's formula and the Euler-Lagrange equations.

Equation (1.6) is the fundamental formula relating the variations of the end values of the state variables with the variations of the control variables. Note that the formula shows the terminal value of the adjoint variable to be the sensitivity of the terminal value of the corresponding state variable to a change in control. This interpretation will be discussed further in connection with the transversal conditions (sect. 1. 9).

Equation (1.6) is often called Green's formula (one form). See Coddington and Levinson (ref. 4) page 86. The last term on

the right applies only at points of discontinuity of F, and since these are discontinuities in $\overset{\circ}{S}$, they are "corners". Hereafter, it will be assumed that the solution does not have a corner unless it is otherwise mentioned. In this case the last term drops out. The fundamental lemma of the calculus of variations states that, given a control program P(t) such that the constraints E are satisfied, then if the coefficient of δP in (1.6) does not vanish for some solution U to the adjoint it is possible to construct a variation such that the constraints are still satisfied, but the function M exceeds the value on the first path. A proof of this lemma is included in refs. (6) and (10). Thus for M to have at least a stationary value M*,

(1.7) $$UF_p = 0. \qquad \text{(The Euler equation)}$$

Equation (1.5) defines the system of equations which is adjoint to the variations of the equations of motion, (1.1). The adjoint equations and equation (1.7) will be called the Euler-Lagrange equations. It will be shown that these, together with the constraints and the equations resulting from the transversal conditions comprise a complete set which determines the optimum solution.

1.6 The maximum principle of Pontryagin.

The functions f_i are often composed of terms which are functions of the state variables alone and terms which contain also the control variables. Suppose we collect terms of the first type

into an n x 1 matrix we will call V, and those of the second type into an n x 1 matrix G. Then (1.1) becomes

$$\dot{S} = V + G.$$

Now consider these matrices to be vectors in an n-dimensional hyperspace of the state variables, and a particularly fruitful geometric interpretation of (1.7) appears. Let us consider equation (1.7)

$$\vec{U} \cdot \vec{F}_p = 0$$

as the condition that

$$\vec{U} \cdot \vec{G} = \max_p$$

that is, let us require that the vector $\vec{G}$ be everywhere so directed that the projection of $\vec{G}$ upon $\vec{U}$ is maximized. The control program which maintains this relationship (while satisfying the constraints) is optimum at least locally. This statement of the condition is due to the Russian mathematician Pontryagin, and it leads to particularly enlightening geometric formulations in many problems. Note that the shape of the domain of $\vec{G}$ then gives insight into the behavior of the control program.

In figure 1.2 a hypothetical domain of $\vec{G}$ is shown by the dashed line. The scalar quantity $\vec{U} \cdot \vec{G}$ is maximized when $\vec{G}$ has a maximum projection on $\vec{U}$ at all times. Figure 1.3 shows the situation where more than one $\vec{G}$ vector satisfies the maximum condition. This generally yields a discontinuity of the control, hence a "corner".

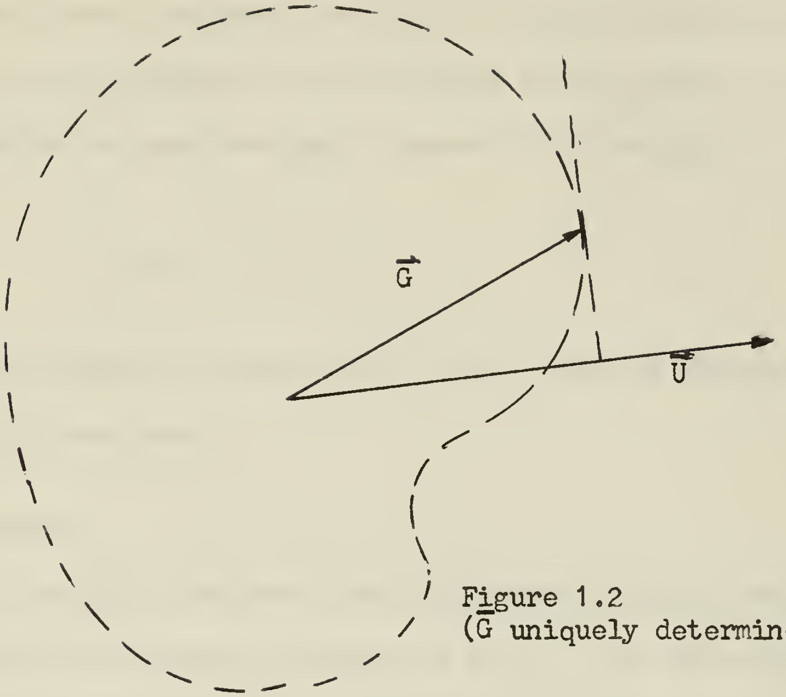


Figure 1.2
($\bar{G}$ uniquely determined)

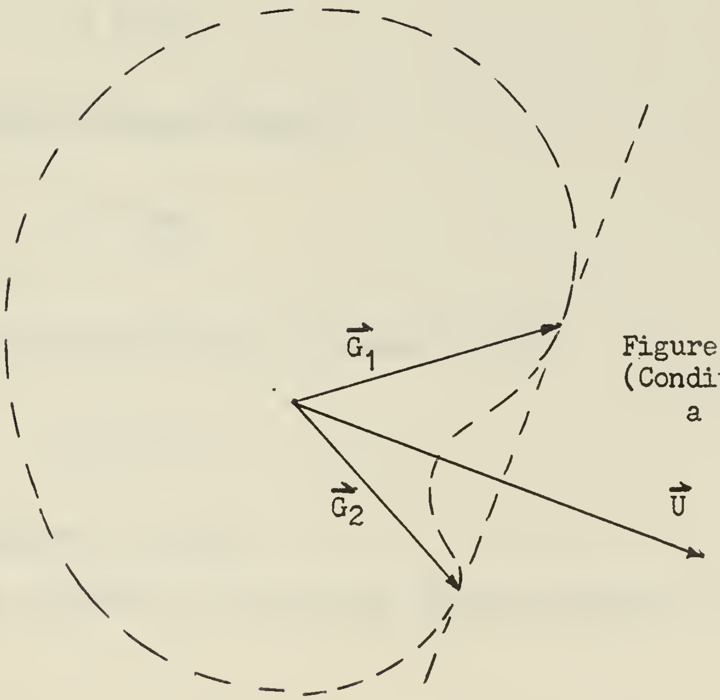


Figure 1.3
(Condition for
a corner)

The functions g_i will be called the "forcing functions". It is the vector whose components are these functions which must be directed so as to maximize its projection on the adjoint vector. The equation or equations which express this condition in the form

$$p_i = h(U, S, P, t)$$

are obtained by ordinary calculus from (1.7). They are sometimes called "steering equations".

1.7 The Hamiltonian.

From the foregoing, one sees that the product UF has special significance when U is chosen according to (1.5). For convenience, we shall define this quantity as the Hamiltonian function, H.

$$H = UF.$$

The maximum principle becomes

$$H = \max_{p}, \tag{1.8}$$

and the general condition for an extreme of H is

$$H_p = 0, \tag{1.9}$$

which is equivalent to (1.7).

For any one extremal, if t does not appear explicitly in the

equations of motion, the Hamiltonian is constant. This may be shown readily, for

$$\begin{aligned} \dot{H} &= \dot{U}F + U\dot{F} \\ &= \dot{U}F + UF_s\dot{S} + UF_p\dot{P} + UF_t \\ &= (\dot{U} + UF_s)F + UF_p\dot{P} + UF_t. \end{aligned}$$

But $\dot{U} + UF_s = 0$ by equation (1.5), and $UF_p = 0$ by (1.7), hence

$$\dot{H} = UF_t = H_t \tag{1.10}$$

In a great many problems, the independent variable does not appear explicitly in the equations of motion, and the Hamiltonian becomes a constant of the system. In the next section it will be shown that in this case if the final value of the independent variable, T, is free (that is, it is not the quantity to be optimized and is not constrained) then H = 0 everywhere on a given path, provided the Euler-Lagrange equations (1.5) and (1.8) are everywhere satisfied on the path.

1.8 Systems which are linear in the state variables.

Consider the functional

$$\int_0^T UF\,dt = \int_0^T H\,dt = \varphi(T).$$

If the differential equations of motion are linear in the state variables, then the functional does not involve the state variables,

and an absolute maximum may be obtained. If the differential equations are not linear, then there are no corresponding general proofs, and we can look only for a stationary value and a local extreme for the functional.

A system of differential equations which is linear in the state variables with constant coefficients may be written as

$$\dot{S} = A\,S + Q$$

where S is the nxl matrix of equation (1.1), Q is an nxl matrix whose ith component is the forcing function of the control variable in the ith direction, and A is an nxn matrix of constants. Now, in analogy with the derivation of equation (1.6), we may form the integral

$$\int_0^T U(\dot{S} - AS - Q)dt.$$

Integrating by parts

$$\Big[US\Big]_0^T = \int_0^T \Big[(\dot{U} + UA)S + UQ\Big]dt.$$

Taking U to be the solution to the adjoint equation $\dot{U} + UA = 0$, we have the form of the Green-Lagrange formula

$$\Big[US\Big]_0^T = \int_0^T UQdt. \tag{1.11}$$

Suppose that we have by some means discovered a set of initial values for the adjoint variables and integrated forward until, at some time T,

H 1. the curve satisfies the constraints on the state variables and the control variable,

H 2. the forcing function Q is such that UQ is maximum everywhere on the curve, and

H 3. the terminal value of the kth component of the adjoint vector is unity, and all other components are zero at $t = T$.

In this case, Faulkner (ref. 5) shows that the curve furnishes an absolute maximum for the kth state variable under the given initial values and the constraints. Akheizer (ref. 6, sect. 2.15) and Edelbaum (ref. 7, sects. 1.1 and 1.2) discuss conditions for strong and weak extremes in this and in nonlinear cases.

1.9 The transversal condition.

The extremals have been discussed, as the solution curves which satisfy the system

$$(1.12)\qquad \begin{cases} \dot{S} - F = 0 \\ \dot{U} + UF_s = 0 \\ H_p = UF_p = o \end{cases}$$

Consider the system already discussed, and the corresponding Green-Lagrange variational equation (1.6),

$$\Big[U\delta S\Big]_o^T = \int_o^T (UF_p\,\delta P)dt - \Sigma_k\Big[UF\Big]_{t_k-}^{t_k+} dt_k = \int_o^T \delta H\ dt.$$

We require that the term on the left, evaluated at the starting point be zero. If $s_i(o)$ is specified, then $\delta s_i(0) = 0$. If s_i is not specified, then we shall choose the particular solution to the adjoint so that $U_i(0) = 0$. It is necessary that the last term on the right be zero, else changing dt_k will offer a better trajectory. There are variations for which δP is not "small" and

$$\int_0^T \delta H \, dt < 0,$$

but for first order effects, the only condition on δS for an extremal is that

$$(U \, \delta S)_T = 0.$$

The form of the equation does not depend on the constraints, hence we will define as admissible paths those which satisfy the constraints. Since the constraints must be satisfied at $t = T$, the endpoint $S(T)$ must lie on the $n - k$ dimensional surface

$$E(S,T) = 0$$

Hence, for an admissible differential at the endpoint,

$$de_i = \left[\sum_j \partial e_i/\partial s_j ds_j + \partial e_i/\partial t dT \right]_T = 0$$

or in matrix notation,

$$(1.13) \qquad dE = (E_s dS + E_t dT)_T = 0$$

where dE and E_t are of dimension kxl, E_s is kxn and dS is nxl. E_s may be considered to be the gradient of E in n-space. If E_s is bordered on the right by the column E_t, this may similarly be considered to be the gradient of E in the n+1 space (of the state variables and t) which we will call ∇E^*.

$$\nabla E^* = \begin{bmatrix} \partial e_1/\partial s_1 & \cdots & \partial e_1/\partial s_n & \partial e_1/\partial t \\ \vdots & & \vdots & \vdots \\ \partial e_k/\partial s_1 & \cdots & \partial e_1/\partial s_n & \partial e_k/\partial t \end{bmatrix}$$

Now if dS^* is taken to be

$$dS^* = \begin{bmatrix} ds_1 \\ \vdots \\ ds_n \\ dT \end{bmatrix}$$

(1.13) may be rewritten as

$$dE = (\nabla E^* dS^*)_T = 0.$$

In the same manner, we may have the gradient of M in n+1 space,

$$\nabla M^* = \begin{bmatrix} \partial M/\partial s_1 & \partial M/\partial s_2 & \cdots & \partial M/\partial s_n & \partial M/\partial t \end{bmatrix}$$

and the augmented adjoint vector

$$U^* = \begin{bmatrix} u_1 & u_2 & \cdots & u_n & (-H) \end{bmatrix}$$

The vector U* must lie in the manifold spanned by the k+1 vectors ∇e_k, ∇M, hence it may be expressed

$$U^* = C\begin{bmatrix}\nabla E^*\\ \nabla M^*\end{bmatrix} = C\,Z^*$$

where C is a $1x(k+1)$ row vector of constants and Z* is the $(k+1) \times (n+1)$ matrix obtained by bordering ∇E^* below with ∇M^*. Since M is assumed to be independent, Z* has rank k+1. The transversal condition may be stated as the condition that the matrices ∇E^*, Z* and Y*

$$Y^* = \begin{bmatrix}Z^*\\ U^*\end{bmatrix} \text{, a } (k+2) \times (n+1) \text{ matrix}$$

all have rank k+1.

As a simple example, suppose we have the state variables w, x, y and z, with all initial conditions specified and for t=T w constrained, x to be maximized, y, T and z free. We may write

w	x	y	z	T	
1	0	0	0	0	constraints
0	1	0	0	0	maximum
u_1	u_2	u_3	u_4	-H	adjoint variables

a 3x5 matrix which must be of rank 2 at $t = T$. Hence

$$u_3 = u_4 = -H = 0 \text{ at } t = T.$$

In the development above, free use was made of material from a research paper of Faulkner.[11] When each e_i and M involves only one variable of the set (S,t), the condition may be summarized as follows:

$$(1.14) \qquad \begin{bmatrix} s_i(T) \text{ free implies } u_i(T) = 0 \\ T \text{ free implies } H(T) = 0 \end{bmatrix}$$

Chapter II

SOLUTION OF A GENERAL OPTIMUM PROGRAMMING PROBLEM USING DIFFERENTIAL CORRECTIONS

2.1 The missing constants of the adjoint set.

It has been shown that if relations (1.12) and (1.14) are satisfied, then the resulting path generally furnishes an optimum to the state S(T). Finding this particular path depends upon finding a particular set of constants, which are the initial values of the unknown adjoint variables. This being true, suppose for the moment that the means exist for correcting a set which is "pretty good", until the final "perfect" set is obtained. That is, we visualize in the domain of U(0), the adjoint vector evaluated at t=0, the point U*(0) whose coordinates are these "perfect" starting values. If U(0) = U*, then integration of the first two of equations (1.11) with the control chosen according to the third of equations (1.12) will produce the optimum path. At some t=T on this path, the constraints will be satisfied, the desired variable will be maximized, and simultaneously (1.12) and (1.14) will be satisfied.

In Chapter III a method will be given that will allow us to correct from an arbitrary point U'(0) toward U*(0) provided U'(0) is within some unknown region R in the vicinity of U*(0). The difficulty is to find some point U'(0) which lies within R. The so-called Direct Method of solution used here reduces the problem to

the solution of a two-point boundary-value problem involving the extremals. Consequently, the missing boundary conditions must somehow be guessed. In general, the guess must be within some region R, if the corrective program is to converge. In the Gradient, or Steepest Ascent methods of Kelley and Bryson a nominal trajectory is guessed and the two-point boundary-value problem is solved by correcting along paths which are not extremals. This method too has its disadvantages. For some problems the optimum path may never be reached, although the correction program "converges", that is, satisfies the convergence criterion. For the problem treated in this thesis, a path was obtained by the direct methods given here which attained a final value of the variables to be maximized well beyond that of the "optimum" path obtained by the method of Steepest Ascent.

2.2 Nominal paths.

An advantage of the Gradient Method is the following. In physical optimization problems we often have some idea beforehand what kind of control program is likely to come fairly close to the optimum one. The only "guessing" involved in the Gradient Method is to choose such a program and construct a nominal path which connects the given initial constraints with the given terminal constraints. Provided the guess is good enough, the path is corrected by making the integrated value of H_p smaller and smaller along the path.

In a given problem, the degree of "goodness" required of this nominal path when using the Gradient Method corresponds in a rough way to the "goodness" of the first guess for $U'(0)$ when using extremals. In the former case, however, one can be guided by physical reasoning. The physical meaning of the value of $U'(0)$ is by no means as clear.

In the preliminary work associated with this thesis, the authors have worked on a means to generate a point $U'(0)$ within R from a nominal path. In the sample problem undertaken, the method works well and provides a value for $U'(0)$ which is quite close to U^*. The method is quite simple and straightforward and adds very little complexity to the computer program which is used to correct from $U'(0)$ to $U^*(0)$.

2.3 The fundamental adjoint set

As a primary tool, we make use of the fundamental set which is used by Faulkner (5). The fundamental set is the nxn matrix,

(2.1)

$$\underline{U} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{bmatrix}$$

whose value at t=0 is the nxn identity matrix and whose elements are chosen so that u_{ij} is the j'th component of the i'th solution

to the adjoint equation (1.5). Note that each row vector of the fundamental set is a particular solution to the adjoint (1.5). For instance,

(2.2) $$\underline{U}_2(0) = \begin{bmatrix} u_{21} & u_{22} & \cdots & u_{2n} \end{bmatrix}_{t=0} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Now, consider a particular combination of the row vectors U_i,

(2.3) $$U = c_1 U_1 + c_2 U_2 + \cdots + c_n U_n,$$

where the c's are constants.

This may be written as

(2.4) $$U = C\underline{U}, \text{ where}$$

(2.5) $$C = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}.$$

Note that as thus defined, $U(0) = C$, so that the C's are the initial values of the adjoint variables. At any time t, u_i is the ith component of U:

(2.6) $$u_i = c_1 u_{1i} + c_2 u_{2i} + \cdots + c_n u_{ni},$$

or

(2.7) $$u_i = C\underline{u}_i, \text{ where}$$

$\underline{u}_i$ is the ith column of the fundamental set.

One of the c's may be chosen arbitrarily. If some state variable

does not appear explicitly in the equations of motion, then the corresponding adjoint variable is constant. If this state variable is to be maximized then it will be convenient to make the corresponding c equal to plus one. If the variable is free at either end point then the corresponding adjoint variable is zero everywhere, and a different c will be chosen arbitrarily. In any case, we must be sure that the corresponding adjoint variable is non-zero.

2.4 Program for generating the C-vector from a nominal path

In this section a program is given for generating a starting set of constants U'. A nominal control program is chosen from physical reasoning which appears likely to be fairly close to optimal and which will generate a path from the given initial conditions to a terminal point which satisfies or nearly satisfies the terminal constraints. This may be simple or difficult, depending on the problem. If such a program cannot be found, then this method for starting cannot be used, nor can be Gradient Method, since both are identical up to this point. In this case, we must go back to the "guessing game". It may be necessary to "map" the U(0) hyperplane to find which regions give good starting points, and then to try various points within those regions. That is, guess the c's and calculate the corresponding trajectories in some systematic way.

Suppose however that a satisfactory nominal control program is found. The equations of motion are now integrated numerically from the given initial conditions using the nominal control program, stopping when a terminal constraint is met. At the same time, we integrate also the n^2 equations of the fundamental set. At the terminal point S(T), each of the elements of the fundamental set has some value $u_{ij}(T)$.

If the path had been the optimum path, equations (1.14) would have been satisfied at the point S(T). We will choose the vector C so that they are satisfied, using the values from the nominal path. Then this vector C becomes U'(0) for starting the corrective iterations. If the program does not converge, the nominal path may be varied somewhat to get another trial C. In this way, the problem of guessing the initial values of the adjoint variables may be eliminated, as in the Gradient Method and still the optimum path is approached through extremals.

It is to be noted that equations (1.14) are not the only relationships which are used to furnish the required number of equations in the C's. According to the particular problem, the Hamiltonian may be constant, or a particular adjoint variable may be constant. These allow use of relations calculated for t=0 rather than at t=T as above.

2.5 Faulkner's method of differential corrections.

A method was given above for obtaining a first approximation for the starting values of the adjoint variables from a nominal path. Let us now suppose that such a set is in hand which was obtained in this way or in some other way.

If the equations of motion (1.1) together with the adjoint equations (1.5) are now integrated numerically with the control chosen according to the maximum principle or (1.7) a path results which is an extremal in that it furnishes an optimum to some end state. But the path obtained using this particular set of initial conditions does not, in general, satisfy the terminal constraints (1.2) and the transversal conditions (1.14). We must have some means for correcting these constants (which we have variously called the C-vector or the vector U(0) in earlier sections) so that (1.2) and (1.14) are satisfied at some point S(T).

The method used here is the method of differential corrections developed by Faulkner (ref. 5) which makes use of differentials in an iterative scheme which is of the Newton-Raphson type. Making use of the fundamental set and the development of Chapter I, the method may be presented quite simply.

Equation (1.6) provides the relation (for a problem without corners)

$$\Big[U\delta S\Big]_0^T = \int_0^T UF_p \delta P dt.$$

For every state variable s_i which is specified at $t = 0$, $\delta s_i(0) = 0$. For those which are free, we will require $u_i(0) = 0$, so that $u_i \delta s_i = 0$ for all cases. This procedure takes care of separated end conditions in a simple and straightforward manner. This case causes a great deal of difficulty with some methods of solution. Note that there are still n unknowns at the start, since the value of s_i is unknown but the value of the corresponding u_i is known.

With this simplification, we have

$$(2.8) \qquad [U\,\delta S]_T = \int_o^T (UF_p)\,\delta P\,dt,$$

where the dimensions are nxn for U, nx1 for δS, nxm for F_p and mx1 for P. From (1.7) we will choose the p's so that

$$H_p = UF_p = 0.$$

Now if S and t are fixed and the c's are varied by small amounts δC, then P must vary so that

$$(2.9) \qquad \delta C U_c F_p + \delta P^T H_{pp} = 0,$$

where the superscript T indicates the transposed matrix. But since $U = C\underline{U}$, then $U_c = \underline{U}$. Making this substitution, and noting also that H_{pp} is the symmetrical mxm matrix whose ij'th element is

$$(H_{pp})_{ij} = \partial^2 H / \partial p_i\, \partial p_j\ ,$$

we now form the mx1 transposed of (2.9).

$$(2.10) \qquad (\underline{U}F_p)^T\, \delta C^T + H_{pp}\, \delta P = 0$$

This equation is solved for δP, and the result is substituted into (2.8).

$$\delta P = -(H_{pp})^{-1}(\underline{U}F_p)^T \delta c^T$$

(2.11) $$\left[\underline{U}\ \delta s\right]_T = -\int_0^T (UF_p)\,(H_{pp})^{-1}\,(\underline{U}F_p)^T\,\delta c^T\,dt,$$

a form convenient for machine programming.

Now suppose each term in the integral on the right be denoted as I_{ji}. We have then at $t = T$

(2.12) $$\begin{bmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & & \vdots \\ u_{n1} & \cdots & u_{nn} \end{bmatrix} \begin{bmatrix} \delta s_1 \\ \vdots \\ \delta s_n \end{bmatrix} = - \begin{bmatrix} I_{11} & \cdots & I_{1n} \\ \vdots & & \vdots \\ I_{1n} & \cdots & I_{nn} \end{bmatrix} \begin{bmatrix} \delta c_1 \\ \vdots \\ \delta c_n \end{bmatrix} .$$

Since one c is arbitrary, say c_1, then $\delta c_1 = 0$ and the equation furnishes n-1 relations among the δc's and the δs's.

In addition, the transversal conditions provide n-k equations whose variational form is, for the problem to be considered here,

$$\delta u_i = \delta C \underline{u}_i .$$

Now for corrections we let

(2.13) $$\delta u_i = -u_i(T) - \dot{u}_i(T)\,\delta T = \delta C \underline{u}_i$$

and also, for the k constrained variables,

$$\delta s_i = -s_i(T) - \dot{s}_i(T)\ \delta T + s_{ci}(T)$$

where $s_{c_i}(T)$ is the given terminal constraint for the ith variable. Since there are k of these variables constrained, there remain (n-k-1) elements δs_i in equation (2.12). These may be eliminated, yielding k equations. The n-k equations (2.13) are adjoined to these, for a total of n equations involving the unknowns δT, δC_2, δC_3, . . . δC_n in terms of the integral elements I_{ij}, the elements of $\underline{U}(T)$, the deriviatives at t=T and the differences (2.13). The equations are solved for the unknowns, and these are applied as corrections for the start of the next iteration.

Chapter III

DEVELOPMENT OF THE PROBLEM

3.1 General description of the problem and assumptions.

The problem described in this thesis is that of determining the optimal angle of attack program p(t), which will maximize the distance covered over the earth's surface by a hypersonic glider or lifting vehicle which has been injected into some initial re-entry point. The starting point of the trajectory is specified by the re-entry injection parameters, the initial conditions, which are:

(3.1)

$v(0) = 25{,}920$ feet/second	-	initial velocity
$h(0) = 300{,}000$ feet	-	initial altitude
$x(0) = 0$ nautical miles	-	initial distance
$z(0) = 0.18$ degrees	-	initial flight path angle

These parameters were taken from reference (8) so that a comparison could be made between the Gradient Method used by earlier authors and the method of differential corrections which was used in this thesis. Also, the value of wingloading, the value of acceleration due to gravity, the value for a standard earth radius, and the value of a standard nautical mile were taken from the same reference.

Wingloading $= \frac{mg_0}{A} = 27.3$ lb/ft^2 , where:

m = mass of the vehicle

A = wing plan form area

$g_0 = 32.2$ ft/sec^2 - acceleration due to gravity at the earth's surface

The model atmosphere used was based on the tables of reference (9). The atmospheric density was assumed to have the form

$\rho = \rho_0 e^{-bh}$, where the parameter b was calculated to fit the tabular values at h = 0 feet and h = 200,000 feet. It was further assumed that the atmosphere was spherically symmetric and fixed with respect to the earth for simplifying reasons.

3.2 Equations of motion.

The coordinate system used is depicted on Fig. 3.1 and the equations of motion are:

$$
(3.2)\quad
\begin{aligned}
\dot{x} &= \frac{R}{R+h} v \cos z, \\
\dot{h} &= v \sin z, \\
\dot{v} &= -\frac{D}{m} - g \cos z, \\
\dot{z} &= \frac{L}{mv} + \frac{v}{R+h} \cos z - \frac{g}{v} \cos z
\end{aligned}
$$

where: $R = 3440$ nautical miles, the radius of the earth,

$$D = \frac{1}{2}(\rho v^2 A\ C_L),$$

$$L = \frac{1}{2}(\rho v^2 A\ C_D),$$

$$g = g_0 \left[\frac{R}{R+h}\right]^2.$$

The coefficients of lift and drag are from reference (8) and are:

$$C_L = C_{LO} \sin p \cos p \ |\sin p|$$

$$C_D = C_{LL} \ |\sin^3 p| + C_{DO}$$

Where: $C_{LO} = 1.82$; $C_{DL} = 1.46$; $C_{DO} = 0.042$; and p is the angle of attack of the vehicle.

3.3 Adjoint equations.

The system of differential equations which is adjoint to the variational equations of the equations of motion (3.2) is:

$$\dot{u}_x = 0$$

$$\dot{u}_h = \left[\frac{Rv}{(R+h)^2}\cos z\right]u_x + \left[\frac{1}{m}\frac{\partial D}{\partial h} - \frac{\partial g}{\partial h}\sin z\right]u_v$$

$$+ \left[\frac{v\cos z}{(R+h)^2} - \frac{1}{mv}\frac{\partial L}{\partial h} + \frac{\partial g}{\partial h}\frac{\cos z}{v}\right]u_z$$

$$\dot{u}_v = \left[\frac{-R\cos z}{R+h}\right]u_x - \left[\sin z\right]u_h + \left[\frac{1}{m}\frac{\partial D}{\partial v}\right]u_v \tag{3.3}$$

$$+ \left[\frac{L}{mv^2} - \frac{\partial L}{\partial v}\frac{1}{mv} - \frac{\cos z}{R+h} - \frac{g\cos z}{v^2}\right]u_z$$

$$\dot{u}_z = \left[\frac{Rv\sin z}{R+h}\right]u_x - \left[v\cos z\right]u_h + \left[g\cos z\right]u_v$$

$$+ \left[\frac{v\sin z}{R+n} - \frac{g\sin z}{v}\right]u_z$$

The fundamental set of adjoint equations $\underline{U}$ or equivalently (u_{ij}) was chosen such that the fundamental set at time $t = 0$ is the identity matrix of rank four. That is:

$$\underline{U} = (u_{ij}) = \begin{bmatrix} u_x^1 & u_x^2 & u_x^3 & u_x^4 \\ u_h^1 & u_h^2 & u_h^3 & u_h^4 \\ u_v^1 & u_v^2 & u_v^3 & u_v^4 \\ u_z^1 & u_z^2 & u_z^3 & u_z^4 \end{bmatrix} \tag{3.4}$$

and further: $(u_{ij})\big|_0^T = \int_0^T(\dot{u}_{ij})dt$, where the $\dot{u}_{ij}$ are defined by the adjoint equations (3.3). We have the further relation that: $\vec{U} = [u_x\ u_h\ u_v\ u_z] = [c_1\ c_2\ c_3\ c_4][\underline{U}]^T$ which is a more convenient form for computations using Fortran programming.

In general, the solution of the adjoint equations is only determined to within a multiplicative constant , hence we can choose one of the c's. In this problem, we chose $c_1=1$.

If we consider the equations of motion in vector or matrix form, we have the following relationships:

$$\vec{S} = \begin{bmatrix} x \\ h \\ v \\ z \end{bmatrix} \quad ; \quad \vec{F} = \dot{\vec{S}} \quad ; \quad \vec{F}_p = \begin{bmatrix} 0 \\ 0 \\ -\frac{D_p}{m} \\ \frac{L_p}{mv} \end{bmatrix} \quad \text{where } \vec{F}_p = \frac{\partial \vec{F}}{\partial p}$$

The Hamiltonian is given by $H = \vec{U}\cdot\vec{F}$.

3.4 Maximum principle.

For any extremal, we must satisfy the condition that $\vec{U}\cdot\vec{F}_p = 0$ for any point on the extremal. This condition uniquely determines the angle of attack program for that particular extremal since:

$$\vec{U}\cdot\vec{F}_p = \frac{\partial L}{\partial p}u_z - v\frac{\partial D}{\partial p}u_v = 0$$

Substituting for $\frac{\partial L}{\partial p}$ and $\frac{\partial D}{\partial p}$ and solving for p we obtain:

$$(3.5) \qquad p = \arctan\left[\left[\frac{3C_{DL}}{2C_{LO}}\,\frac{v}{u_z}\right]\left[\sqrt{[u_v]^2 + \frac{8}{9}\left|\frac{C_{LO}}{C_{DL}}u_z\right|^2} - u_v\right]\right]$$

$$\text{for } -\frac{\pi}{2} < p < \frac{\pi}{2}$$

3.5 End conditions.

Since we wish to maximize the total distance traveled over the earth's surface, we must have the following conditions at the end time $t = T$:

$x(T) = \text{maximum}$

(3.6) $h(T) = 0$

$v(T)$, $z(T)$, and T free.

3.6 Variational equations.

On any curve without corners, we have the following relation between the adjoint and the variations:

$$(3.7) \qquad [u_x^j\delta x + u_h^j\delta h + u_v^j\delta v + u_z^j\delta z]_T = \int_0^T \vec{U}^j\cdot\vec{F}\ \delta p\ dt \qquad j = 1,2,3,4$$

For an extremal we must also satisfy the Euler equations or

the maximum principle of Pontryagin so that we have $\int_0^T \vec{U}\cdot\vec{F}\,dt$ = maximum and $\vec{U}\cdot\vec{F}_p = 0$ along the extremal. If we consider x,h,v,z, and t fixed and change the constants then $\left[\frac{\partial\vec{U}}{\partial c_2}\delta c_2 + \frac{\partial\vec{U}}{\partial c_3}\delta c_3 + \frac{\partial\vec{U}}{\partial c_4}\delta c_4\right]\cdot\vec{F}_p + \vec{U}\cdot\vec{F}_{pp}\delta p = 0$. Noting that $\frac{\partial\vec{U}}{\partial c_i} = \vec{U}^i$, i =2,3,4, we can solve for δp and substitute in equation (3.7) which will give us four equations with which we can correct the initial values of the c's for a terminal variation of the altitude h:

$$(3.8)\qquad \left[u_x^j\delta x + u_h^j\delta h + u_v^j\delta v + u_z^j\delta z\right]_T = -\sum_{i=2}^{4}\int_0^T \frac{(\vec{U}^j\cdot\vec{F}_p)(\vec{U}^i\cdot\vec{F}_p)}{\vec{U}\cdot\vec{F}_{pp}}dt\,\delta c_i$$

Let $I_{ij}=\int_0^T \frac{(\vec{U}^j\cdot\vec{F}_p)(\vec{U}^i\cdot\vec{F}_p)}{\vec{U}\cdot\vec{F}_{pp}}dt\,\delta c_i$

Since $u_x^1 = 1$, $u_x^2 = u_x^3 = u_x^4 = 0$ for all t, and noting that $I_{ij} = I_{ji}$, we can solve the following set of equations for δh in terms of the δc_i:

$$(3.9)\qquad \begin{bmatrix} u_h^2 & u_v^2 & u_z^2 \\ u_h^3 & u_v^3 & u_z^3 \\ u_h^4 & u_v^4 & u_z^4 \end{bmatrix}_T \begin{bmatrix} \delta h \\ \delta v \\ \delta z \end{bmatrix} = \begin{bmatrix} I_{22} & I_{23} & I_{24} \\ I_{23} & I_{33} & I_{34} \\ I_{24} & I_{34} & I_{44} \end{bmatrix}_T \begin{bmatrix} \delta c_2 \\ \delta c_3 \\ \delta c_4 \end{bmatrix}$$

From the transversal conditions, we have that:

$$(3.10)\qquad \begin{aligned} &u_v(T) = 0 \\ &u_z(T) = 0 \\ &(\dot{x}u_x+\dot{h}u_h)_T = 0 \end{aligned}$$

From the last equation of (3.10) we have: $u_h(T) = (-\cot z)_T$

Since the Hamiltonian $H = \vec{U}\cdot\vec{F}$ and $H(T) = 0$ from equations (3.10), then the Hamiltonian is identically zero for all t. This fact will be used to determine one of the constants (c_2) and effectively reduce the problem of guessing the initial c's to start the calculations.

The transversal conditions (3.10) give us three more independent variational equations associated with the end conditions:

$$\delta u_v(T) = [u_v^2\delta c_2 + u_v^3\delta c_3 + u_v^4\delta c_4]_T$$

$$(3.11)\quad \delta u_z(T) = [u_z^2\delta c_2 + u_z^3\delta c_3 + u_z^4\delta c_4]_T$$

$$\delta[\dot{x}u_x + \dot{h}u_h]_T = \dot{h}(T)[u_h^2\delta c_2 + u_h^3\delta c_3 + u_h^4\delta c_4]_T$$

3.7 Differential corrections

The method of differential corrections given by Faulkner in reference (5) was used for the solution of the optimization problem. In general, we will guess the initial values associated with the adjoint equations and the terminal time T. For this problem, we must choose c_2, c_3, c_4, and T. We then correct these parameters by calculating an extremal with the initial guess and then solve for corrections to the parameters in terms of the end values associated with the extremal. For all end conditions of the form $s(T) = S$, we will generally have some error in $s(T)$. We set $\delta s(T) = S - s(T) - \dot{s}(T)\,\delta T$. For this problem, we set $\delta h(T) = h(T) - \dot{h}(T)\delta T$ and in a similar manner, we obtain from equation (3.11):

$$-u_v(T) = [u_v^2\delta c_2 + u_v^3\delta c_3 + u_v^4\delta c_4]_T + \dot{u}_v(T)\delta T$$

$$(3.12)\quad -u_z(T) = [u_z^2\delta c_2 + u_z^3\delta c_3 + u_z^4\delta c_4]_T + \dot{u}_z(T)\delta T$$

$$[\cot z]_T = \dot{h}(T)[u_h^2\delta c_2 + u_h^3\delta c_3 + u_h^4\delta c_4]_T$$

$$+[\ddot{x}u_x + \ddot{h}u_h + \dot{h}\dot{u}_h]_T\delta T$$

These equations together with the solution to (3.9) for δh in terms of the δc's will give a set of equations from which we

solve for the δc's and δT to make corrections in the starting values for the subsequent trajectory.

3.8 Computation of approximate starting values by use of nominal trajectories.

In general, for a nth order set of differential equations, one must choose (n-1) of the constants of the solution to the adjoint equations to start the problem. This is generally a difficult problem in itself, since the constants are not related to the physical aspects of the problem in any discernable manner. One of the primary aspects of this thesis was to find a method which could be used to find a logical choice for the initial valres of these constants. Fortunately the Hamiltonian was identically zero in this problem. This fact, together with the transversal conditions (3.10) enabled us to solve for approximate constants by first computing a nominal trajectory; here we choose a nominal angle of attack program p(t), which seems likely in a physical sense to give a maximum distance trajectory. The first such nominal angle of attack program tried was to use the maximum lift over drag ratio where p=20.5 degrees, a constant. This did not work. We then used the following angle of attack program for which we are indebted to Professor Faulkner:

$$p(t) = 54.7^\circ \left[1 - \frac{h(t)}{750{,}000} \right]$$

The value of 54.7° for p is the value at which we have maxium lift. It was supposed that the angle of attack was never negative and from the data of constant angle of attack programs, the value of 750,000 feet was chosen to insure that p was never negative.

With this choice for the angle of attack program, we obtained a good set of starting constants by integrating the fundamental set $\underline{U}$ along with the equations of motion until $h = 0$. At this time, we solve the set of equations:

$$(3.13)\quad \begin{bmatrix} u_h^2 & u_h^3 & u_h^4 \\ u_v^2 & u_v^3 & u_v^4 \\ u_z^2 & u_z^3 & u_z^4 \end{bmatrix}_T \begin{bmatrix} c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} -u_h^1 - \cot z \\ -u_v^1 \\ -u_z^1 \end{bmatrix}_T$$

for c_2, c_3, and c_4. These equations are the transversal equations (3.10). This problem of guessing the c's is further reduced due to the Hamiltonian being identically zero, since for a given choice of c_3 and c_4, the angle of attack is uniquely determined from equation (3.5) and then we may solve for c_2 from the Hamiltonian at time $t = 0$. If we have the case where $\dot{h}(0) = 0$, then we have a singularity and connot solve for c_2, but in this case, the choice of c_2 is arbitrary. Effectively, this problem is then reduced by an order of one and we need only guess two constants. We felt that if we had an extremal, we would determine its constants in this manner and hence we could get estimates for the constants from a trajectory which was a "good" approximation to an extremal.

3.9 Computational scheme.

This problem was programmed on a CDC 1604 computer using a Runge-Kutta integration method. The computational scheme is as follows:

i) We need a nominal trajectory to start. To get it, we first intergrate the equations of motion (3.2) and the fundamental set (3.4) using a programmed angle of attack and ending the

trajectory when $h = 0$. At this time, we solve for the initial constants c_2, c_3, and c_4 from equation (3.13). We then use these constants to start the first optimal trajectory.

ii) Then we compute the first optimal trajectory where the angle of attack program is determined by the maximum principle. We integrate the equations of motion, the fundamental set, and the nine I_{ij} in equation (3.9) a total of twenty-five equations. We terminated the integration when one of the following conditions is met: $h = 0$, $u_v = 0$, $u_h = 0$, or $|z| \geqq \frac{\pi}{2}$, since it was felt that the end of the significant part of the trajectory had been reached. We felt that in this problem no optimum would occur for $u_v < 0$ since the resulting high drag would dissipate energy needlessly. The supposition that this is a maximum distance trajectory dictates that $|z| < \frac{\pi}{2}$.

Since we do not have t appearing explicitly in the equations and since the terminal conditions determine T, we do not use δT in our calculations. We only need the δc's the correct the c's to start the next trajectory.

iii) We then repeat the computations as in the above paragraph, using the new c's for the next trajectory. We test for convergence at the end of each trajectory:

(3.14) If $[h(T)]^2 + [\dot{x}u_x + \dot{h}u_h]_T^2 + [u_v(T)]^2 + [u_z(T)]^2 \leqq \epsilon$,

we say that we have converged to the solution.

3.10 Results and conclusions.

The optimum trajectory that was calculated is depicted on Fig.3.3 and the corresponding angle of attack program is depicted on Fig. 3.4. On these two figures, a comparison has been made between the results of this thesis and results from reference (8), which were obtained by the Gradient (or Steep-

est Ascent) Method. The comparison can not be considered as an exact one since the integration method, the time step or integration interval, and the approximation to the atmospheric models may not be identical. However, there appears to be a significant increase in the maximum distance obtained using the method of differential corrections.

The convergence criteron (3.14) we initially chose turned out to be a poor choice. We could not determine how close we were to the optimal solution. We then proceeded to "map" the plane of c_3 and c_4 by using a systematic choice of c_3 and c_4 and computing the corresponding extremals. The results of the mapping procedure are given on Fig.3.5, Fig.3. ,and Fig.3.7. Fig.3.5 shows the contours of distances obtained by extremals as a function of c_3 and c_4. Fig.3.6 gives detailed contours around the optimal solution. Fig.3.7 is a cross section of the "ridge" which occurs in the c_3 and c_4 plane. It is a conjecture of the authors that some difficulties are encountered by a gradient method solution when there exists such a ridge in a problem, since the slope is nearly zero.

The major problem encountered was the apparent instability of the problem near the end of an extremal. We attempted to integrate backwards from the terminal conditons and found that this was virtually impossible. The problem seems to be unstable; we could not even integrate backwards with constant p. Backward integration was very good if we started backward from a point that was not too close to the end of the trajectory.

Another main problem encountered was that of determining

the end time on a trajectory. We first thought that all we need consider is the time when $h = 0$. However, we found that other conditions effectively terminated the useful part of the extremal earlier. For example, most extremals led to a point where $|z| > \frac{\pi}{2}$.

Theoretically, on the maximum distance trajectory, all terminal conditions would be met simultaneously. In the actual computations, this did not occur. The most probable reason that we have this apparent discrepancy is that the round off errors in the calculations and the inherent error in the integration method are large enough to prevent us from satisfying all conditions at the end time.

In the final analysis, we found that a program which reduced the magnitude of the corrections would lead to convergence. The method used was to calculate a set of c's from a nominal trajectory and then calculate the first trajectory and the δc's corresponding to that trajectory. We store x, the c's, and δc's of this trajectory and then calculate the next trajectory. If the distance of this trajectory is greater than the distance of the preceding one, we proceed with the iterative process. If the distance is less, then we reduce the corrections by one half and calculate the trajectory again. If we still have less distance, we further reduce the corrections until we obtain a greater distance. The reason we developed this method was because the magnitudes of the δc's were of the order of magnitudes of the c's. It was noted that the integrals relating δh to the δc's are improper integrals. This, in addition to the unstable char-

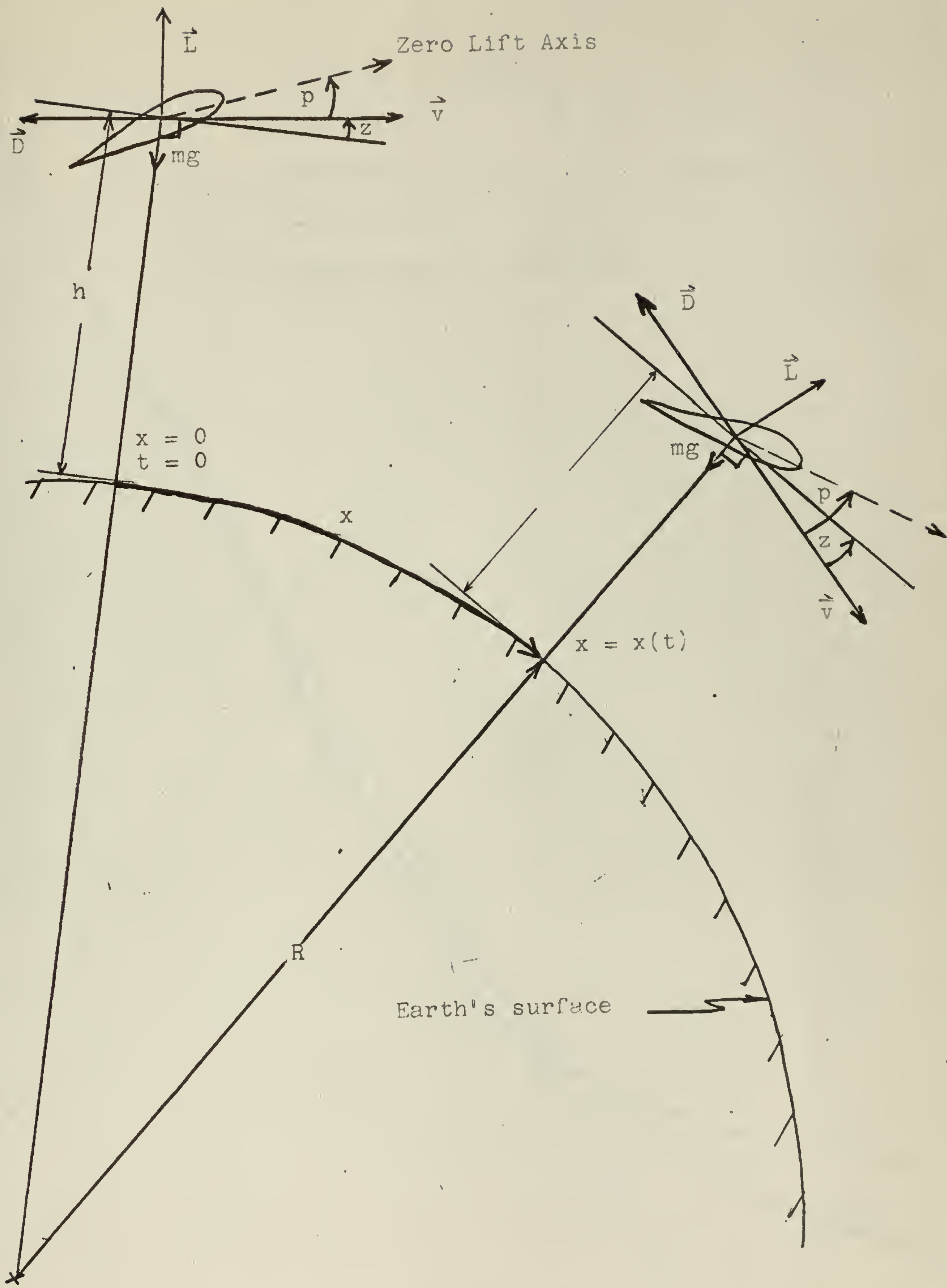
acter of the end of the trajectory, tended to make the magnitude of the δc's large; however, the program did correct in the right direction. These integrals seem to be divergent for the extremals which satisfy the transversal conditions. A formulation of the problem suggested by Professor Faulkner which would avoid these divergent integrals by using a set of u's which were determined at the end time was programmed. However, this program was not checked out or run since the original integrals did not actually diverge in calculations made on the computer. The method suggested was to use a fundamental set $\underline{U}^*$ such that $\underline{U}^*(T) = \delta_{ij}$. After computing an extremal using $\underline{U}$, we solve for the matrix of constants $[a_{ij}]$ which satifies the relation: $\underline{U}^*(T) = [a_{ij}]\underline{U}(T)$. We then compute the same extremal using $\vec{U}^* = \vec{U}_1^* + c\vec{U}_2^*$, where $\vec{U}_i^* = \sum_{j=1}^{4} a_{ij}\vec{U}^j$, i= 1,2. We then have the relation

$$\delta h(T) = -\int_0^T \frac{(\overline{U}_2^* \cdot \overline{F}_p)^2}{\overline{U}^* \cdot \overline{F}_{pp}} dt\ \delta c$$

and we solve for δc to correct the starting values. Further, this integral is not divergent.

The first integration method used in calculations was rectangular integration and it was found that trajectories thus calculated were very rough. A comparison of the rectangular integration method and the Runge-Kutta method was made and we chose the latter method due to its accuracy and smoothness.

Figure 3.1



Coordinate System

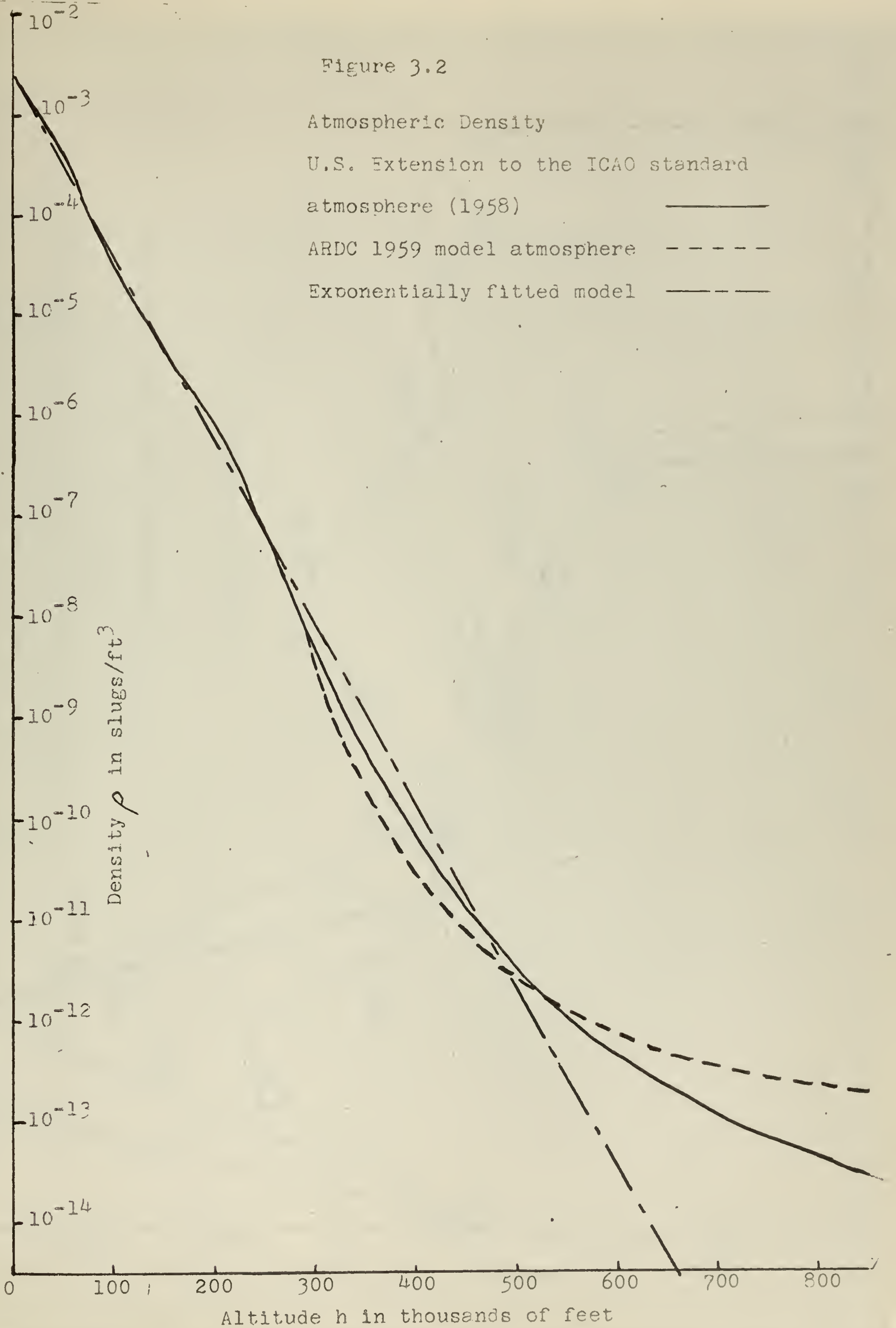
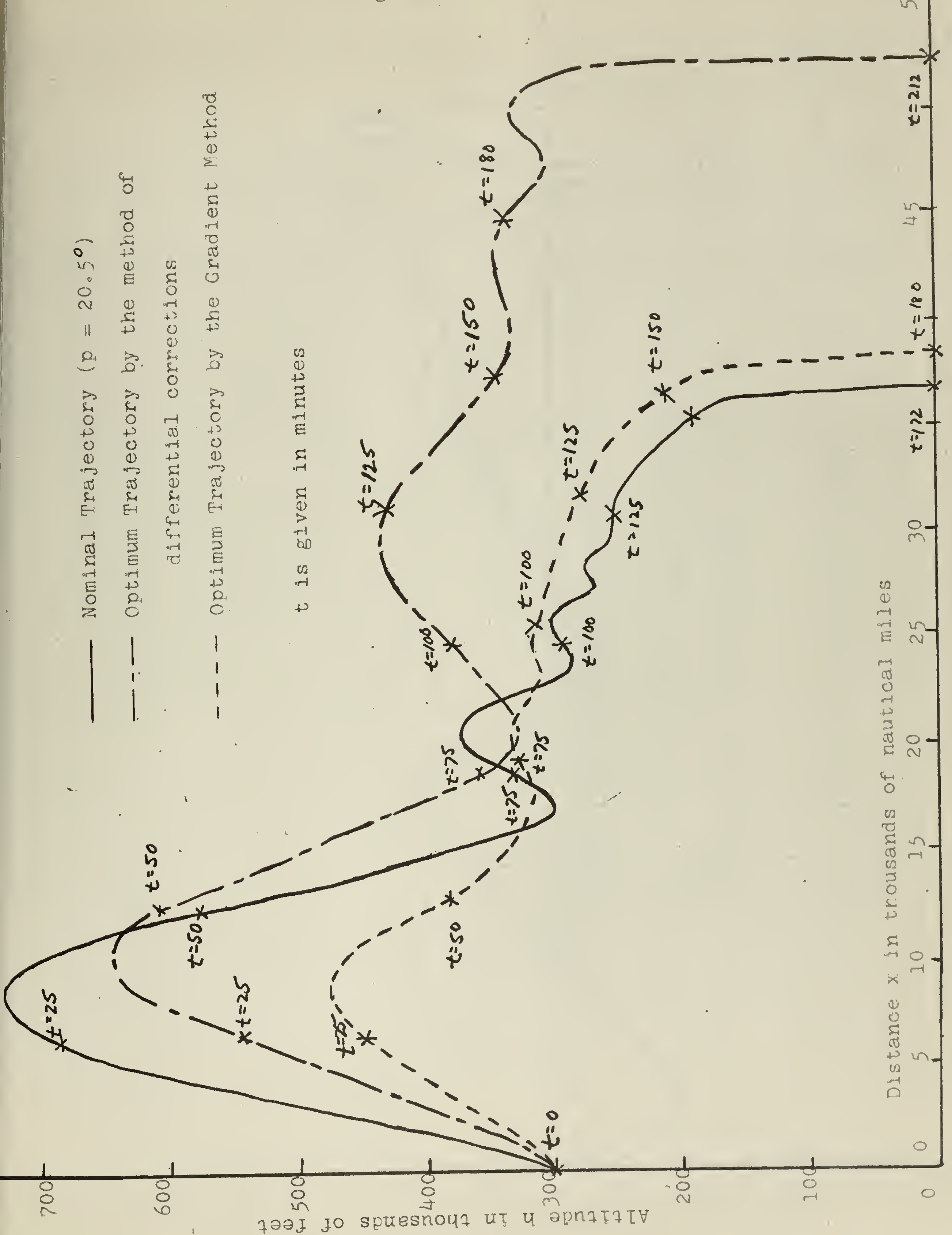


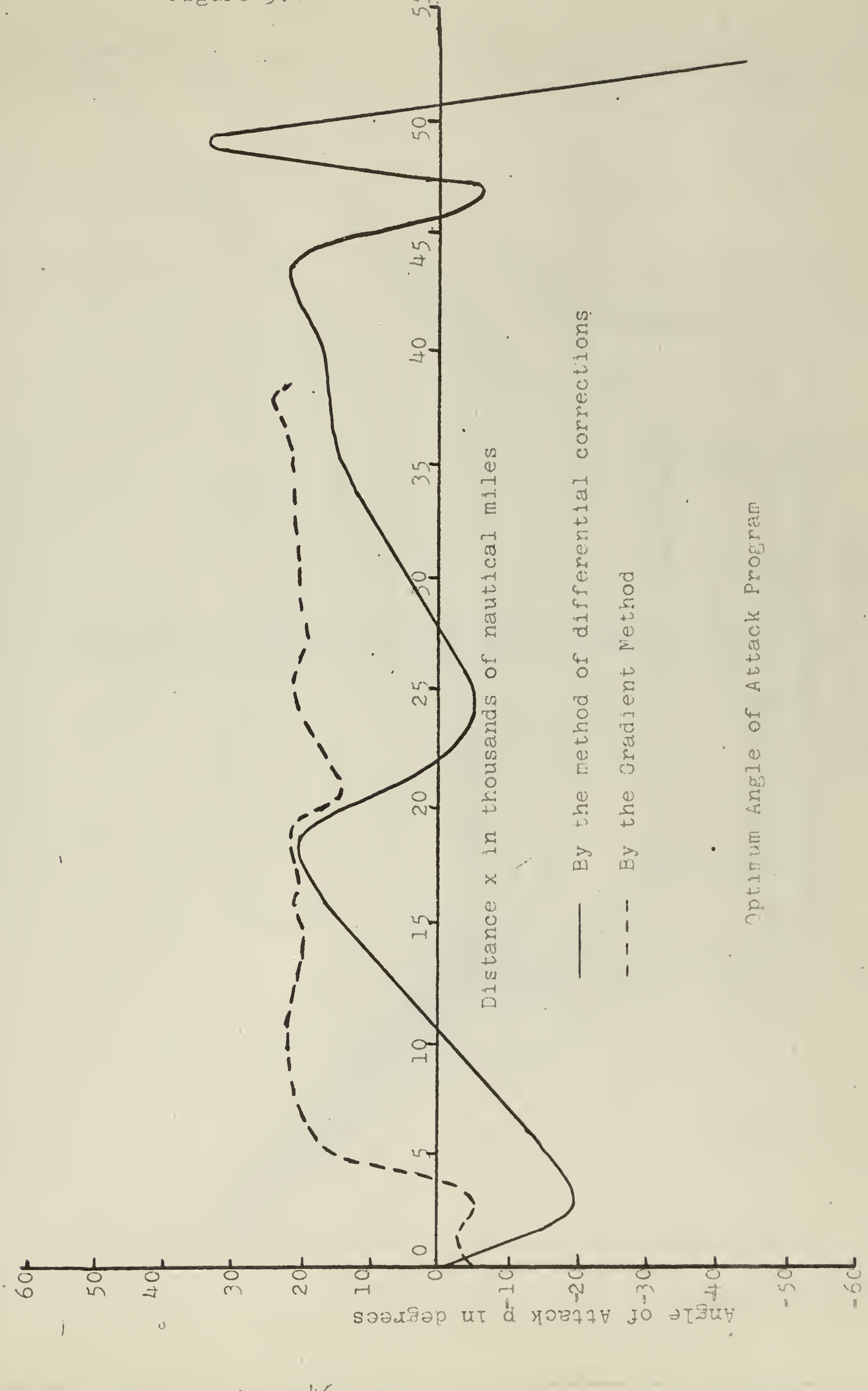
Figure 3.2
Atmospheric Density
U.S. Extension to the ICAO standard atmosphere (1958)
ARDC 1959 model atmosphere
Exponentially fitted model
Density ρ in slugs/ft^3
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
10^{-10}
10^{-11}
10^{-12}
10^{-13}
10^{-14}
0
100
200
300
400
500
600
700
800
Altitude h in thousands of feet

Figure 3.3



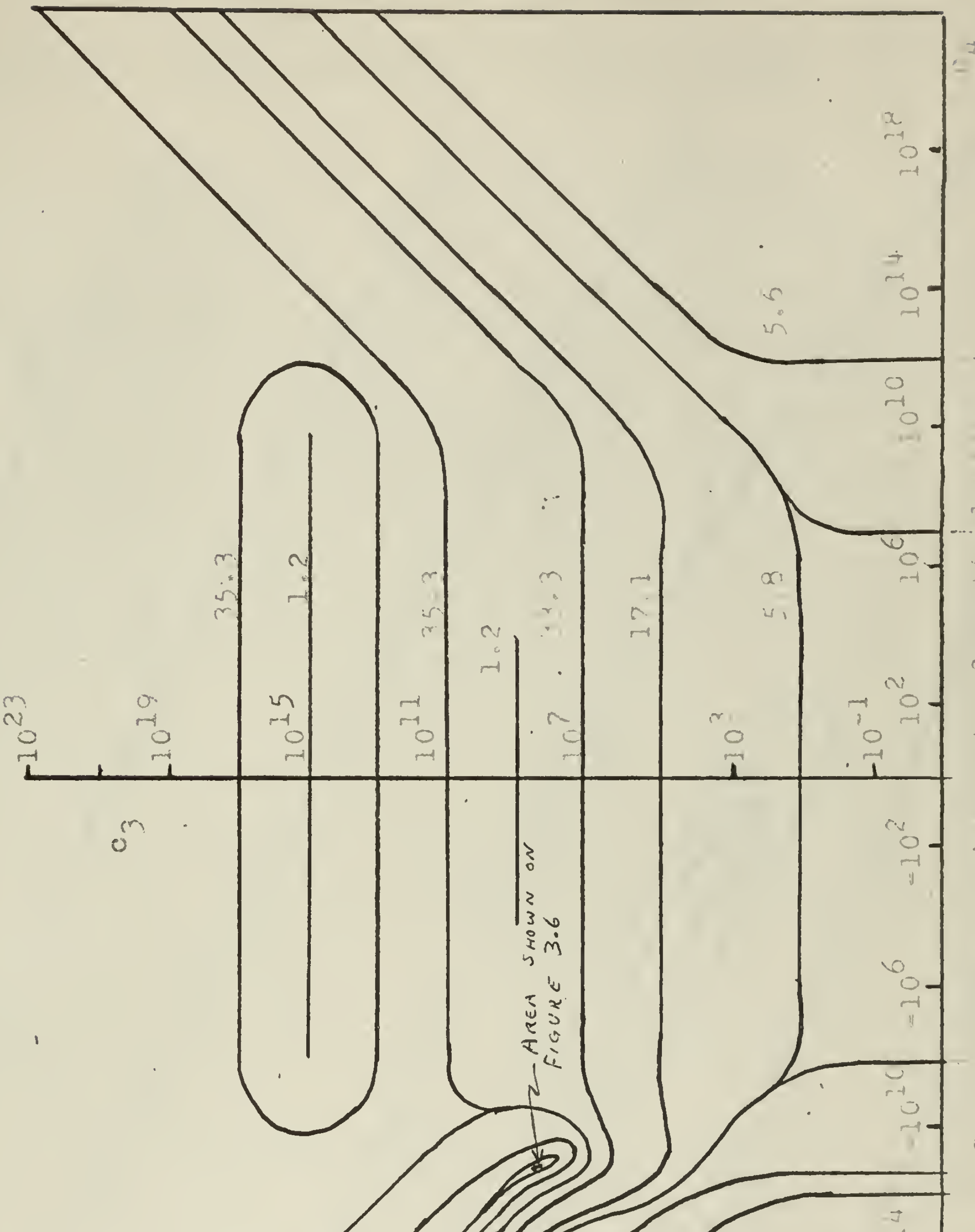
Nominal Trajectory (p = 20.5°)
Optimum Trajectory by the method of differential corrections
Optimum Trajectory by the Gradient Method
t is given in minutes
Distance x in trousands of nautical miles
Altitude h in thousands of feet

Figure 3.4



Distance x in thousands of nautical miles
0
5
10
15
20
25
30
35
40
45
50
55
Angle of Attack p in degrees
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
By the method of differential corrections
By the Gradient Method
Optimum Angle of Attack Program

Figure 3.5



Contours are in thousands of nautical miles

Map of $c_3 - c_4$ plane

Figure 3.6

Detailed Contours of the c_3 - c_4 plane

Contours are given in thousands of nautical miles

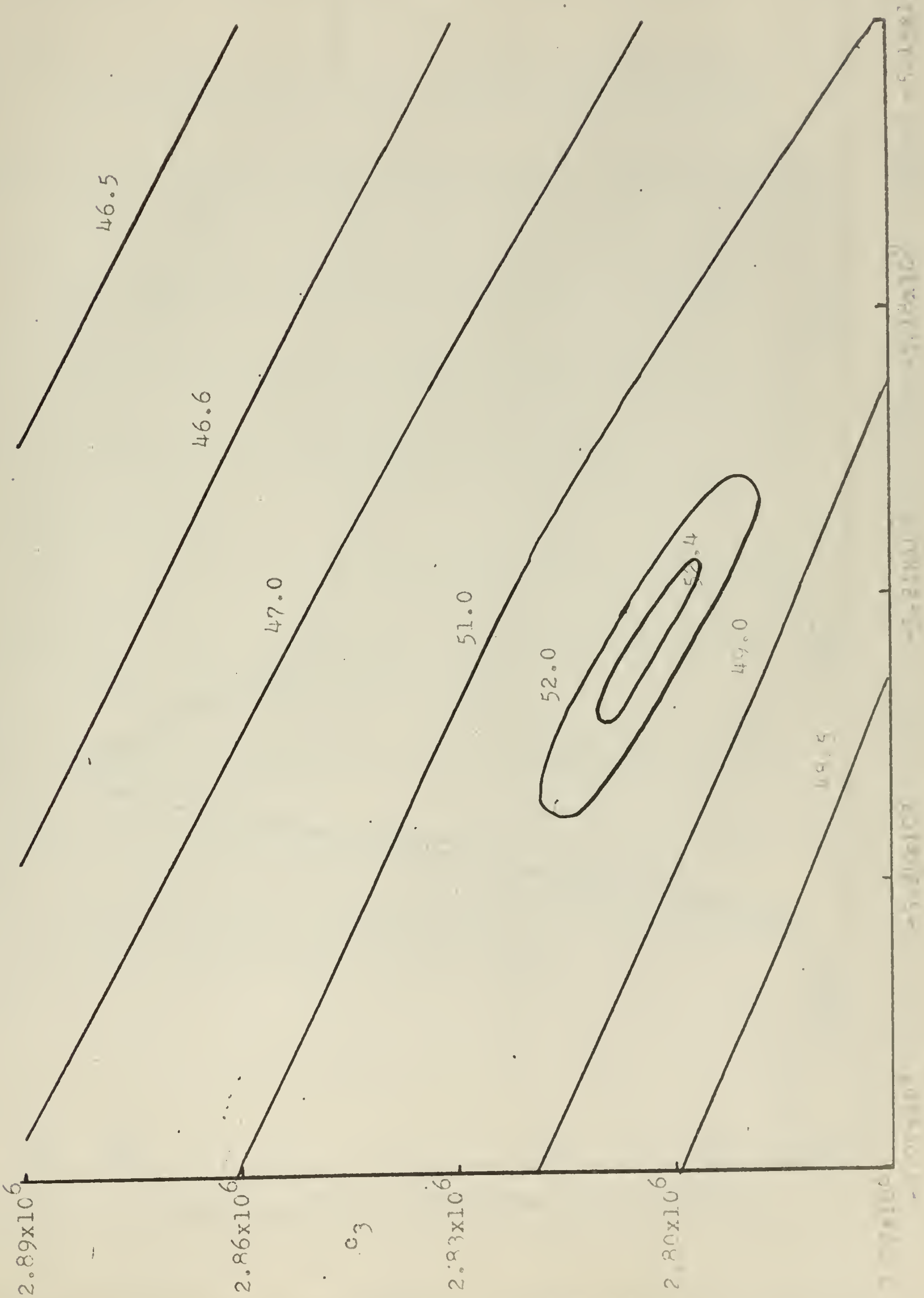
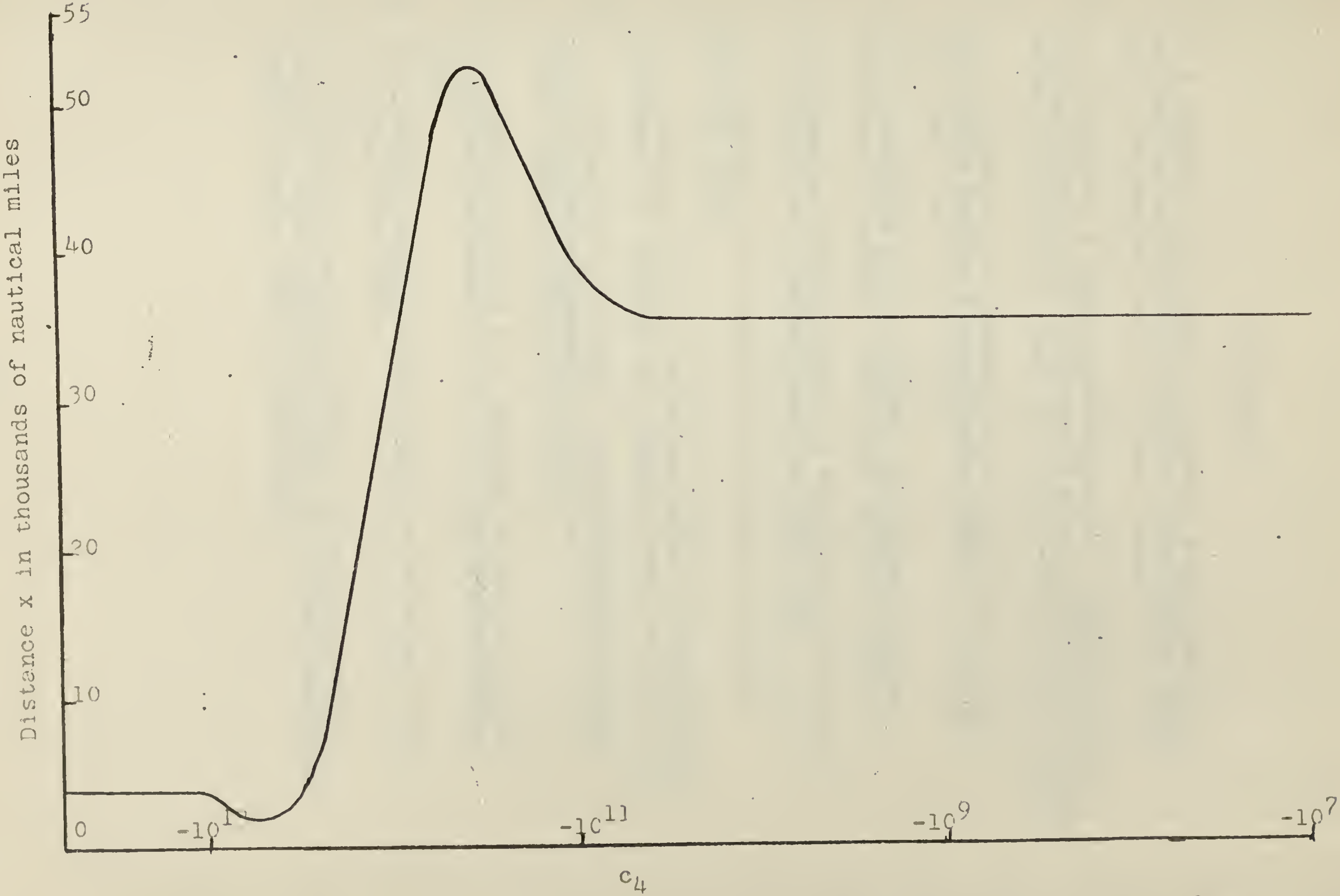


Figure 3.7



Cross Section of Contours of c_3 - c_4 plane along the line: $\log_{10}c_3 = \log_{10}c_4 + 19$

BIBLIOGRAPHY

1. F. A. Valentine, The problem of Lagrange with differential inequalities as added side conditions, Dissertation, Department of Mathematics, Univ. of Chicago, Chicago, Ill., 1937.

2. F. D. Faulkner, A simple interpretation of a degenerate Euler equation arising in the ship-routing problem, Presented at meeting of No. Calif. Sect. Am. Math. Soc., 1952 (Available from author).

3. J. V. Breakwell, The optimization of trajectories, Rept. AL-2706, North American Aviation Company (1957); J. Soc. Ind. Appl. Math. Vol. 7 no. 2 (1959).

4. E. A. Coddington and N. Levinson, "The Theory of Ordinary Differential Equations", McGraw-Hill, New York, 1955.

5. F. D. Faulkner, Direct methods, in "Optimization Techniques", G. Leitmann, ed., Academic Press, New York, 1962.

6. N. I. Akheizer, "The Calculus of Variations", Blaisdell, New York, 1962.

7. T. N. Edelbaum, Theory of maxima and minima, in "Optimization Techniques", G. Leitmann, ed., Academic Press, New York, 1962.

8. A. E. Bryson and W. F. Denham, A steepest-ascent method for solving optimum programming problems, Ratheon Rpt. BR-1303, Ratheon Company, Bedford, Mass., 1961.

9. R. A. Minzer, (et al.) U. S. extension to the ICAO standard atmosphere, Washington, D. C., U. S. Govt.Printing Office, 1958.

10. G. A. Bliss, "Lectures on the Calculus of Variations", Univ. of Chicago Press, 1946.

11. F. D. Faulkner, A method of solution for optimum control problems, USNPGS Research Paper #36, U. S. Naval Postgraduate School, Monterey, Calif., (Forthcoming).